A DOCUMENTAL DATABASE OF CHINESE SEISMOLOGY JOURNAL PAPER (CSJP)

Gao Shuxin

(*The Earthquake Research Institute of Lanzhou, SSB*)

Abstract

This paper describes the CSJP database comprehensively, which has been developed by the Information Research Office, the Earthquake Research Institute of Lanzhou. It is designed by VAX DATATRIEVE, VAX Common Data Dictionary softwares on VAX 11/750 minicomputer, so the CSJP database is an English documental database. It mainly collects journal papers about the earthquake prediction research from 50 plus journalswhich are edited and issued in China. Among of them, 16 journals are edited and issued by seismological institutes or provincial seismological bureaus which belong to State Seismological Bureau. Referring the ISO and Chinese biliographical description rules, we design the format and data items [fields] of a documental record standard and in common use. The length of a record for a paper is fixed, 1898 bytes long, including abstract 1020 bytes. There are about 800 plus papers every year. Tapes or disks of the documental database can be exchanged or sold at home and abroad.

INTRODUCTION

Since 1976, our country has imported 50 plus foreign documental magnetic tapes. These magnetic tapes have provided convenient conditions for Chinese users to retrieve foreign documents. But, to the Chinese documents, our country are lack of unified plan, organization, and constructing documental databases using modern tools. Now the fact is that it is more difficult to retrieve Chinese documents than to retrieve foreign documets, and so it makes Chinese documents difficult to enter international exchange. Therefore, we are planning to constructuct the CSJP databases [English edition and Chinese edition] for the users at home and abroad, For the first step, we have constructed this

CSJP database [English edition] under present conditions. Here it is described comprehensively.

The CSJP database is a file-structured database set up with VAX DATATRIEVE [DTR] and VAX Common Data Dictionary [CDD] softwares on VAX 11/750 minicomputer [VAX/VMS operation system. Peripheral equipments: TU80 magnetic tape unit, RX02 diskette drive, VT220 and HZ—8401A terminals, CYD—903 printer etc.].It mainly collects scientific journal papers about the earthquake prediction research from 50 plus journals which are edited and issued in China. Among of them, 16 journals are edited and issued by seismological institutes or provincial seismological bureaus which belong to State Seismological Bureau. We use the English biliographical description of every Chinese paper to index and to form the secondary documents, so the CSJP database is an English documental database. The documental amount is about 800 plus every year.

Fig. 1 is the structural diagram of retrieval system of the CSJP database. It's a on-line retrieval system. The general structural design had passed the appraisal of State Seismological Bureau in Aug. 1986.
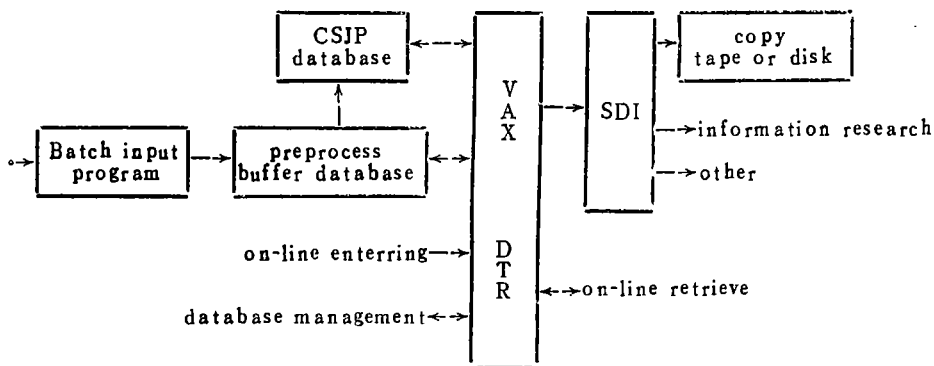

Fig. 1　Structural diagram of retrieval sysatem of CSJP database

SUPPORT SOFTWARES OF CSJP DATABASE

1. VAX DATATRIEVE [DTR]

DTR is a comprehensive data management tool. It operates effectively in commercial, technical, scientific, industrial, and educational environments. DTR is a "fourth-generation language". Its syntax is more English-like than that of BASIC or COBOL. It lets us define data structure and store these definitions separately from any program. Then we can write any number of programs that use the same data structure without having to redescribe the data each time. With DTR, we access the data through constructs called "domain". A dom-

ain definition establishes a name for a set of data and tells DTR where that data is described and where the data is stored. A domain definition contains the name of the domain, the name of a record [data description], and the name of a data file. Now in CSJP database, the name of the domain is CSJP, the name of the record is CSJP-REC, and the name of the data file is CSJP.DAT. Besides the structural design of the database, we have also used DTR to set up many practical procedures, such as retrieval programs which are provided for users to on-line retrieve. Our practice has already proved that DTR is very suitable for documental database of seismological science.

2. VAX Common Data Dictionary [CDD]

The CDD is a VAX software product and is actually a hierarchy dictionaries. DTR uses the CDD to store data definitions and procedures. In an information management system, reliable data definitions are as important as the data itself. Shared data definitions must be unambiguous and sensitive data definitions must be protected. The CDD is assists in these tasks by providing a central storage place for data definitions and a data security system for their protection. With DTR, we set protections and securities for the data definitions and procedures. In operation system VAX/VMS, we use DCL to set protections for our data files.

3. Batch Process Program

Due to using DTR procedures, the speed of on-line entering data is comparatively slow. And also because the length of a record exceeds 255 characters, the data file can not be edited in VAX/VMS. So we have programmed the batch process program using COBOL. Under EDT of VAX/VMS, we enter the original data. And then the edited file can be converted into DTR record format of CSJP database by batch process program. By using this program, the entering speed is increased greatly.

RECORD FORMAT OF CSJP DATABASE

We referred many international and Chinese standards about the biliographical description for the record definition of CSJP database to make the record definition standard and in cnmmon use. The elementary fields [data item] which contain in one record are consistent with the data items which contain in INSPEC retrieval output format of ESA information retrieval service. The record definition is as foll-

```
ows:
RECORD CSJP—REC;1
01 CSJP.
03 CN  PIC  9[6].
03 TI  PIC  X[255] EDIT—STRING IS T[75].
03 AU  PIC  X[70].
03 AA  PIC  X[70].
03 JN  PIC  X[70].
03 YR  PIC  9[4].
03 VN  PIC  X[15].
03 PP  PIC  X[15].
03 CO  PIC  X[6].
03 SN  PIC  X[10].
03 TC  PIC  X[3].
03 LA  PIC  X[2].
03 RF  PIC  9[2] EDIT—STRING IS Z9.
03 AB1 PIC  X[255] EDIT—STRING IS T[75].
03 AB2 PIC  X[255] EDIT—STRING IS T[75].
03 AB3 PIC  X[255] EDIT—STRING IS T[75].
03 AB4 PIC  X[255] EDIT—STRING IS T[75].
03 CC  PIC  X[25].
03 DE  PIC  X[70].
03 ST  PIC  X[255] EDIT—STRING IS T[75].
;
```

Where CN: control number, TI: title of paper, AU: author[s], AA: author affiliation, JN: title of higher level publication, YR: date of publication, VN: volume and issue number, PP: inclusive page numbers, CO: CODEN[code number], SN: ISSN, TC: treatment code, LA: language, RF: numbers of references, CC:classification code, DE: subject index headings, ST: free-indexing terms.

DTR limits the character long in one elementary field. None exceeds 255 characters, therefore one abstract length is divided into four elementary fields [AB1, AB2, AB3, AB4]. As an example, one documental record in CSJP database is as follows:

```
CN    : 860142
TI    :
Several order clustering methods for the study of seismogenic process of strong earthquakes
AU    : Wang Biquan, Chen Zuyin, Tong Guobang, Wang Chunzhen
AA    : Inst. of Geophysics, State Seismological Bur., Beijing, China
JN    : Acta Seismologica Sinica
YR    : 1986
VN    : Vol.8, No.2
PP    : 113—126
CO    : ASSID7
SN    : 0253—3782
TC    : T
LA    : CH
RF    : 14
AB1   :
The seismogenic process of strong earthquakes of North China and the Ningxia region is
studied by using five order clustering methods. It is found that all five methods can simul-
ate the phenomena of alternatively high and low
```

AB2　　：
seismic activities, the better effects of both Fisher's method and order hierarchical cluste-
ring method show that the structures of time intervals in classes are mainly of cluster
distribution rather than chain

AB3　　：
distribution, and it is important that this character of seismic activity should be considered
in choosing the method of prediction. The results of Webster's method and nonparametric
goodness of fit show that there are

AB4　　：
certainly objective valleys between relatively quiet periods and remarkably active periods.
Graph theory further helps to distinguish the smaller peaks in which strong earthquakes
are apt to occur

CC　　　：P315.7 / P315.8 / P315.5
DE　　　：earthquake / geophysical technique / seismology / statistics
ST　　　：
earthquake space-time distribution / geophysical technique / nonparametric fitting method
/ Fisher's method / North China / Ningxia region / order clustering method / strong
earthquake / seismic activity

Using the function of DTR literal connection, the values in AB1,
AB2, AB3, and AB4 can be connected to form a complete abstract.
There are several retrieval output formats, which depend upon the
retrieval programs programmed by ourselves. As an example, a comp-
lete retrieval output format is as follows:

CN　　860142
TI　　Several order clustering methods for the study of seismogenic process of strong ear-
　　　thquakes
AU　　Wang Biquan, Chen Zuyin, Tong Guobang, Wang Chunzhen
AA　　Inst. of Geophysics, State Seismological Bur., Beijing, China
JA　　Acta Seismlogica Sinica
SO　　Vol.8, No.2, PP113—126, 14 Refs, 1986
CODEN: ASSID7　　ISSN 0253—3782.　　TC: T　　LA: CH
AB　　The seismogenic process of strong earthquakrs of North China and the Ningxia regi-
　　　on is studied by using five order clustering methods. It is found that all five methods
　　　can simulate the phenomena of alternatively high and low seismic activities, the bet-
　　　ter effects of both Fisher's method and order hierarchical clustering method show
　　　that the structures of time intervals in classes are mainly of cluster distribution rat-
　　　her than chain distribution, and it is important that this character of seismic acti-
　　　vity should be considered in choosing the method of prediction. The results of Web-
　　　ster's method and nonparametric goodness of fit show that there are certainly
　　　objective valleys between relatively quiet periods and remarkably active periods. Gra-
　　　ph theory further helps to distinguish the smaller peaks in which strong earthquakes
　　　are apt to occur
CC　　P315.7 / P315.8 / P315.5
DE　　earthquake / geophysical technique / seismology / statistics
ST　　earthquake space-time distribution / geophysical techniques / nonparametric fitting
　　　method / Fisher's method / North China / Ningxia region / order clustering
　　　method / strong earthquake / seismio activity

WORKING PROCESS OF CONSTRUCTING CSJP DATABASE

The editoral committe of CSJP database consists of about 20 nu-
mbers and is in charge of indexing, editing and examining the docum-
ental work sheets. Based on the ISO and Chinese biliographical stan-
dards, we have instituted the "Indexing Guidelines of CSJP Database"

and drawn the work sheet of documental indexing. Constructing database adopts the technical line of "unified guideline-dividing up the indexing work-concentrated process". Every indexing number must observe the Indexing Guidelines of CSJP Database in order to guarantee the quality of the database. Fig. 2 is shown the working process of constructing CSJP database. Thesaurus [subject index heading i. e. controlled terms] indexing adopts "Chinese Thesaurus". Free-indexing terms [i. e. uncontrolled terms]are such key words which are extracted from the words used in the title of paper and abstract. Classified indexing [classification code] adopts "Classification for Chinese Books and Reference Materials". We have begun the documental input work since July, 1987. By the end of 1988, there were about three thousand papers [1985, 1986, 1987 etc.] in CSJP database. The documental amount of three years [about 2500] will be entered to the database every year from 1989 to 1990. This on-line retrieval database is planning to store the documental amount of 15 years.
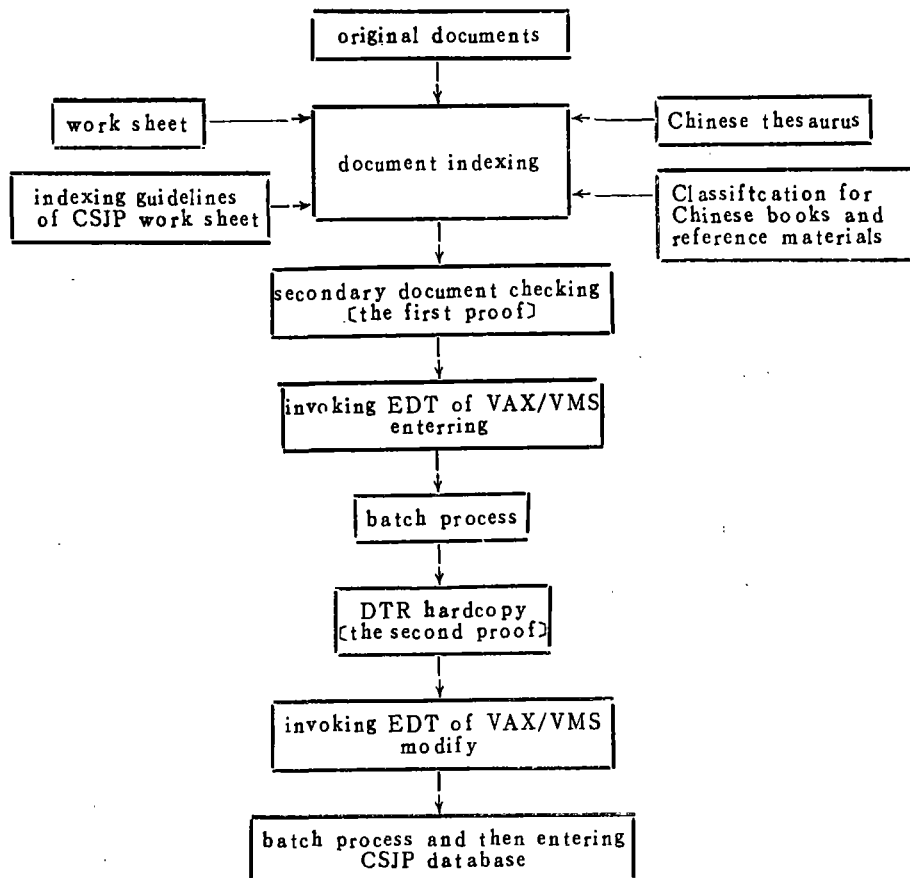

Fig. 2　Working process for setting up CSJP database

CONCLUSION

Although the CSJP database is an English database (secondary documents), most of the original papers are written in Chinese and so it is difficult to refer the original papers for foreign users. If the foreign users retrieve the CSJP database and hope to refer the original papers, we can help them to translate the original papers from Chinese to English. Welcome exchange or buying CSJP tapes or disks.

REFERENCES

[1]VAX DATATRIEVE User's Guide
[2]VAX DATATRIEVE Reference Manual

中国地震科技文献数据库（英文版）

高树心

（国家地震局兰州地震研究所）

摘 要

本文全面地描述了中国地震科技文献数据库（英文版）。该库在VAX 11/750小型计算机上，使用VAX数据检索、VAX公共数据字典软件进行设计。是一个联机检索系统。文献库收集五十余种中国编辑出版的科技期刊中，有关地震预报研究的论文。参考国内外文献著录准则，使文献记录格式标准化，通用化。一条记录长1898字节，其中文摘长1020字节。每年文献量约800余条。文献库磁带和软盘拷贝在国内外交换与出售。